Journal of Organizational Behavior

Journal of Organizational Behavior, J. Organiz. Behav. **37**, 157–167 (2016)
Published online 25 September 2015 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/job.2053

Statistical control in correlational studies: 10 essential recommendations for organizational researchers

THOMAS E. BECKER[1]*, GUCLU ATINC[2], JAMES A. BREAUGH[3], KEVIN D. CARLSON[4], JEFFREY R. EDWARDS[5] AND PAUL E. SPECTOR[6]

[1]*College of Business, University of South Florida Sarasota–Manatee, Sarasota, Florida, U.S.A.*
[2]*Department of Management, Texas A&M University–Commerce, Commerce, Texas, U.S.A.*
[3]*College of Business Administration, University of Missouri–St. Louis, St. Louis, Missouri, U.S.A.*
[4]*Pamplin College of Business, Virginia Tech, Blacksburg, Virginia, U.S.A.*
[5]*Kenan-Flagler Business School, University of North Carolina, Chapel Hill, North Carolina, U.S.A.*
[6]*Department of Psychology, University of South Florida, Tampa, Florida, U.S.A.*

Summary Statistical control is widely used in correlational studies with the intent of providing more accurate estimates of relationships among variables, more conservative tests of hypotheses, or ruling out alternative explanations for empirical findings. However, the use of control variables can produce uninterpretable parameter estimates, erroneous inferences, irreplicable results, and other barriers to scientific progress. As a result, methodologists have provided a great deal of advice regarding the use of statistical control, to the point that researchers might have difficulties sifting through and prioritizing the available suggestions. We integrate and condense this literature into a set of 10 essential recommendations that are generally applicable and which, if followed, would substantially enhance the quality of published organizational research. We provide explanations, qualifications, and examples following each recommendation. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: statistical control; research methods; correlational studies

Organizational researchers often use statistical control in correlational studies, intending to provide more accurate estimates of relationships among predictor and criterion variables, conduct more conservative tests of their hypotheses, or rule out alternative explanations for empirical findings. However, including control variables (CVs) in analyses raises a number of important conceptual and analytical issues. Methodologists have demonstrated that it is important to address these issues because failing to do so can result in uninterpretable parameter estimates, cause inferential errors, hinder replication of results, and in other ways hamper scientific progress (e.g., Burks, 1926; Meehl, 1970; Spector, Zapf, Chen, & Frese, 2000).

Fortunately, recent authors have offered recommendations for properly using CVs in organizational research (Atinc, Simmering, & Kroll, 2012; Becker, 2005; Breaugh, 2008; Carlson & Wu, 2012; Spector & Brannick, 2011). These articles have been cited more than 700 times, and the corresponding recommendations have been incorporated into the norms of the *Academy of Management Journal* and the *Journal of Organizational Behavior* (Bono & McNamara, 2011; Edwards, 2008). Despite this attention, there appears to be little improvement in how organizational researchers handle CVs. For example, Becker (2005) and Breaugh (2008) recommended that authors provide clear justification for including CVs. Nonetheless, Bernerth and Aquinis (in press) reviewed 580 articles containing CVs published in top management journals between 2003 and 2012 and found that in 2003, only 5% of articles included a clear theoretical justification for including CVs. In 2012, the rate was 3%.

*Correspondence to: Thomas E. Becker, College of Business, University of South Florida Sarasota–Manatee, Sarasota, Florida, U.S.A. E-mail: teb1@sar.usf.edu
Excepting the first author, the order of authorship is alphabetical.

 Received 02 September 2015, Accepted 05 September 2015

The literature on the proper use of statistical control has expanded to the point that well-intentioned researchers might have trouble sifting through and prioritizing the recommendations most relevant to their work. Therefore, our objectives are to integrate and condense this literature into a set of 10 essential recommendations that are generally applicable and which, if followed, would substantially enhance the quality of published research. We aim to provide guidance to help organizational scholars apply statistical control appropriately in their research, and, to this end, we state each recommendation concisely and assertively. The discussion following each recommendation provides explanations, qualifications, and examples. We do not supply mathematical formulae or proofs, but, where necessary, we provide citations that do.

Recommendations

Table 1 summarizes our recommendations regarding the treatment of CVs in organizational research. We have organized the recommendations around four issues: selecting CVs, including CVs in hypotheses, measuring and analyzing CVs, and reporting and interpreting results.

Selecting control variables

1. When in doubt, leave them out!

We have borrowed this admonition from Carlson and Wu (2012) because its adoption by researchers would generally improve the interpretation of results. A cautious approach to selecting CVs requires that researchers be certain that statistical control is justified and that a specific set of CVs is appropriate to include in a given study. There are several common situations in which researchers should doubt the value of statistical control.

Table 1. Top 10 recommendations regarding the treatment of control variables in organizational research.

1. When in doubt, leave them out!
Improves the interpretation of results.
2. Select conceptually meaningful CVs and avoid proxies.
Promotes appropriate statistical control and the valid measurement of CVs.
3. When feasible, include CVs in hypotheses and models.
Obviates unjustified inclusion of CVs and fosters more thoughtful hypothesis tests.
4. Clearly justify the measures of CVs and the methods of control.
Discourages proxies and facilitates the interpretation and replication of results.
5. Subject CVs to the same standards of reliability and validity as are applied to other variables.
Fosters construct validity of CVs and increases the accuracy of parameter estimates for the independent variables.
6. If the hypotheses do not include CVs, do not include CVs in the analysis.
Encourages appropriate hypothesis testing and model specification.
7. Conduct comparative tests of relationships between the IVs and CVs.
Contributes to understanding the causal role of CVs and relationships among other variables.
8. Run results with and without the CVs and contrast the findings.
More fully reveals the effects of CVs on the relationships between the independent and dependent variables.
9. Report standard descriptive statistics and correlations for CVs, and the correlations between the measured predictors and their partialled counterparts.
Facilitates understanding of the psychometric properties of CVs, enhances potential for replication, and enables comparison between measured and partialled predictors.
10. Be cautious when generalizing results involving residual variables.
Improves the assessment of external validity and the practical application of results.

DOI: 10.1002/job

Unclear purpose for controlling variables

An unclear purpose for including CVs is likely to cloud the interpretation of results because statistical control changes the meanings of the original substantive variables. With the addition of CVs, parameter estimates (e.g., regression coefficients and path coefficients) reflect not the original measured independent variables (IVs) but the portions of the original IVs that are unrelated to the CVs (Breaugh, 2008; Williams, Vandenberg, & Edwards, 2009). These changes in meaning occur owing to partialling, the mathematical partitioning of variance explained by a set of predictor variables (Nimon & Oswald, 2013; Pedhazur, 1982).

Explained variance can be separated into two broad categories, one that includes variance in the dependent variable (DV) uniquely accounted for by each predictor variable and another that consists of variance jointly explained by two or more predictors. Statistically, joint variance cannot be separated into components attributable to individual predictors but instead become part of the overall variance explained by the set of variables (R^2 in regression analysis). If the predictors are uncorrelated, there is no joint variance, and coefficients for the predictors can be interpreted unequivocally. However, as correlations among predictors increase, so does the proportion of joint variance. Consequently, the amount of joint variance explained by each predictor becomes increasingly ambiguous. This phenomenon is relevant to the use of CVs, because as the correlations between the IVs and CVs increase, the meaning and interpretation of the IVs depart from how the variables were defined and measured prior to analysis. With greater joint variance among IVs and CVs, partialling can make the interpretation of parameter estimates equivocal, particularly when the researcher is uncertain about the purpose of statistical control.

We therefore urge researchers to omit CVs whose purposes are unclear or not conceptually defensible. To ensure a clear purpose, prior to initiating a study, researchers need to consider whether there is a convincing justification for using any type of statistical control (Breaugh, 2008). Examples of when statistical control would not be needed are when all of the variables in a study are of substantive interest (i.e., none are considered to be "nuisance" variables), an area of research is in its early stages, a researchers are only interested in simple correlations among a small number of variables, or prior theory and research are insufficient for justifying CVs.

Uncertain associations between CVs and other variables

Spector and Brannick (2011) demonstrated that the misspecification of relations between CVs and other variables can produce spurious associations among IVs and DVs. Spuriousness occurs when a CV is a cause of multiple IVs, which do not themselves have direct causal connections to the DV. As Burks noted nearly 90 years ago,

> …we are partialling out too much when we render constant factors which may in part or in whole be caused by either of the two factors whose true relationship is to be measured, or by still other unmeasured remote causes which also affect either of the two isolated factors (1926, p. 534).

Even when a CV does not affect the underlying constructs (latent variables), it can cause contamination by influencing the observed measures of the constructs. Thus, unless there is a sound reason for including a CV, adding it can obfuscate rather than clarify relationships among the variables of interest.

Further, we take exception to the common viewpoint that larger numbers of CVs constitute a better, more rigorous methodological approach than including fewer or no CVs. This viewpoint is based on the flawed assumption that adding CVs necessarily produces more conservative tests of hypotheses and reveals the true relations among variables of interest (the fallacious purification principle discussed by Spector and Brannick, 2011). Without knowing the causal relations of variables under consideration, CV usage does not result in more conservative hypothesis tests by "playing it safe." As Meehl (1970) pointed out, "One cannot label a methodological rule as playing it safe when it is likely to produce pseudo-falsifications" (p. 147). In other words, pursuing the goal of conducting conservative tests can result in throwing the baby out with the bathwater (Spector et al., 2000). Also, including large numbers of CVs reduces degrees of freedom, and, even if the CVs are only moderately correlated with the IVs, this will increase standard errors and potentially decrease the power of the test for a given IV.

DOI: 10.1002/job

Impotent CVs

Finally, doubt can be created by what Becker (2005) called impotent CVs, that is, CVs that have little or no relationship with the DV (e.g., $|r| < .10$). He asserted that impotent CVs will not substantively influence results and omitting them will typically increase power and simplify analysis, reporting, and interpretation. We agree that impotent CVs can often be omitted but not necessarily for the reason that Becker suggested. In fact, correlations between the CVs and IVs can affect the results even when the correlation between the CV and DV is zero. However, inclusion of a CV that is unrelated to the DV would not usually satisfy the purposes of statistical control. For example, consider a researcher who hypothesizes that more conscientious employees are tardy to work less often than less conscientious employees. The researcher decides to control for distance from home to work to rule out the alternative explanation that tardiness is simply a function of distance: Employees who live further away from work tend to be tardy more often. If the CV (distance from work) is uncorrelated with the DV (tardiness), then distance cannot be an alternative to conscientiousness as an explanation of tardiness. Similarly, if the correlation between distance and tardiness is approximately zero, it is unclear that including distance in the analysis produces a "more conservative" test of the hypothesized relation between conscientiousness and tardiness. Regarding accuracy of estimates, it is true that if conscientiousness and distance are related (perhaps more conscientious employees buy homes closer to work), then the coefficients for conscientiousness could be significantly different in equations with and without distance. However, the difference in coefficients would not prove or even suggest that the analysis containing distance to work reveals a truer estimate of the effect size of conscientiousness.

In sum, "When in doubt, leave them out!" refers to when CVs should not be used. We turn next to what researchers should do when they have a valid purpose for statistical control.

2. Select conceptually meaningful CVs and avoid proxies.

We urge researchers to thoughtfully choose CVs and corresponding measures because careful selection will promote effective control and valid measurement. The chief goal is to identify specific CVs that match the researcher's purpose (Carlson & Wu, 2012). Toward this end, whenever possible, researchers should provide a strong theoretical rationale for including each CV, and if CVs that theory suggests should be controlled are omitted, then a justification should be offered for their omission (Breaugh, 2008). Selected CVs could represent components of a target theory that are not the primary focus of a study or they might represent alternative theoretical explanations for the hypothesized effects of the focal IVs. Where a strong foundation in theory is not available, researchers should at least provide a logical explanation for selecting a given CV and explain why the CV is a biasing rather than substantive variable (Becker, 2005; Bono & McNamara, 2011). Without these justifications, it is possible that the CV plays a substantive rather than extraneous role in the network of relations the researcher is studying. For example, in the context of job stress research, Spector et al. (2000) discussed the problems that can occur when negative affectivity is routinely treated as a nuisance factor to be statistically controlled. These authors pointed out that if, as some evidence suggests, negative affectivity plays a substantive role in stress-related phenomena, controlling for negative affectivity can lead to removing the effects of the variables (e.g., work conditions) that one wants to study. Researchers should also provide supporting evidence and citations, including a description of the nature of the evidence (e.g., correlational, longitudinal, or experimental) and the expected direction (positive or negative) of associations (Atinc et al., 2012). These requirements will encourage the use of conceptually meaningful CVs.

We distinguish conceptually meaningful CVs from CVs that serve as proxies, surrogates that are used in place of meaningful CVs. Proxies often covary with the DV but are not drawn from theory and do not offer a direct explanation for the association between a meaningful CV and DV. Common examples of proxies are demographic variables, organization size, industry membership, and research setting. A key problem with proxy CVs is that the researcher usually does not know the strength of the relationship between the proxy and the actual CV of interest (Breaugh, 2008). Without this knowledge, one cannot reasonably conclude that the relation between the proxy and other variables is identical or even similar to the relations between the CV of interest and other variables. Hence, researchers testing hypotheses that include proxies can produce different results than would have occurred if a more meaningful CV had been used (Carlson & Herdman, 2012).

Moreover, because the proxy might relate to other variables in a way that the CV of interest does not, controlling for the proxy may control for a host of unintended variables that have substantive effects that the researcher does not wish to remove. For example, a researcher could wish to control for career stage because it is expected to influence the variables of interest. Using age as a proxy for career stage would not only be imprecise but also might control for job tenure, skills that come from experience, and other variables, some of which are unrelated to career stage and that are not intended to be controlled.

In sum, researchers should aim to include conceptually meaningful CVs that are theoretically relevant rather than proxy CVs in their investigations. Bernerth and Aguinis (in press) provide a decision tree composed of sequential steps in the process of selecting CVs, and readers looking for a step-by-step approach may find the decision tree helpful.

Including control variables in hypotheses

3. When feasible, include CVs in hypotheses and models.

We endorse this recommendation because it will help researchers obviate unjustified inclusion of CVs and foster more thoughtful hypothesis tests. Because hypotheses guide research design and analysis, researchers should base their selection of CVs more on theoretical than analytical grounds. Good research practice requires strong logical connections between theory and hypotheses, and this principle applies to CVs just as it does to variables of substantive interest (Carlson & Wu, 2012). The central issue is whether the theory in question mandates the use of CVs to assure that the theory is appropriately tested in a particular context. If so, then CVs should be included in the corresponding hypothesis (Edwards, 2008; Spector & Brannick, 2011). Unfortunately, this practice is rare. For instance, Schjoedt and Bird (2014) examined 140 articles containing CVs published in four prestigious entrepreneurship journals in 2012 and found no studies that included CVs in the hypotheses. Although the practice is uncommon, it is not unheard of, at least in the field of organizational behavior. For example, Loi, Yang, and Diefendorf hypothesized that, “Within individuals, daily interpersonal justice is positively related to daily job satisfaction, after controlling for the effect of daily positive emotions” (2009, p. 772), and Wright and Bonnett hypothesized that,

> Controlling for differences in gender, age, ethnicity, and job performance, PWB [psychological well-being] will moderate the relation between job satisfaction and voluntary turnover such that this association will be weak when PWB is high and negative when PWB is low (2007, p. 147).

We encourage researchers to emulate these examples.

There may be times when a researcher has some evidence supporting inclusion of a CV but is lacking a strong theory-based rationale. Or there may be cases where two or more theories offer competing views regarding inclusion of CVs. In these instances, we endorse Spector and Brannick’s (2011) advice to generate competing hypotheses of the role the CVs might play in relations among the IVs and DVs. Note that none of these circumstances would involve atheoretical CVs, that is, ones for which no explanation is offered. Rather, we are referring to situations in which the strength of the explanation may vary from strong, in which case CVs should be included in the statement of hypotheses, to somewhat tentative, in which case competing hypotheses pertaining to the CVs should be specified.

More broadly, researchers should make explicit in their models the presumed causal structure by which CVs relate to one another and to the substantive variables (Edwards, 2008). Including CVs in models encourages researchers to provide a sound explanation for the role the CVs plays. It also helps to ensure that the analyses match the hypotheses.

Measuring and analyzing CVs

4. Clearly justify the measures of CVs and the methods of control.

Cogently describing measures of CVs is essential for the interpretation and replication of results. Further, the need to explain how each CV was measured and why it was measured that way should discourage the use of proxy CVs.

DOI: 10.1002/job

Regarding methods of control, there are two situations in which the explanation of methodology deserves special attention. The first is whenever an uncommon method of statistical control is used. For example, researchers using structural equation modeling (SEM) often neglect to discuss which paths to and from CVs were freed and fixed. Furthermore, authors routinely omit CVs from path diagrams reporting SEM results. This omission makes it difficult for readers to accurately interpret the findings and for other researchers to replicate the analyses. Per Recommendation 2, the paths from CVs to endogenous variables and correlations with exogenous variables should be specified based on relevant theory, or at least on sound reasoning and empirical evidence.

The second situation deserving special attention is when certain CVs are included in some analyses but not others or some CVs are treated differently than others in the same analysis. For instance, one method of control involves incorporating a CV into the DV by dividing scores on the DV by scores on the CV. A hypothetical example is a researcher dividing sales performance (the DV) by the number of customers in sales territories, thereby producing an index of sales performance controlling for territory. When this approach is adopted, the researcher should explain why it was used instead of more conventional methods (e.g., including number of customers in the regression equation). The need for explanation is most acute when one or more CVs are controlled in this manner while other CVs are controlled through other means.

5. Subject CVs to the same standards of reliability and validity as are applied to other variables.

We argue that researchers should take the psychometric properties of CVs as seriously as they do their IVs and DVs. Adopting this mindset will encourage proper attention to the construct validity of CVs and potentially increase the accuracy of parameter estimates for IVs. An important step in this direction is for researchers to choose measures of CVs that correspond to the concepts that CVs are intended to represent. Becker (2005) discussed an example of when correspondence was not convincingly demonstrated. The study in question measured nursing home size, the CV, as the natural log of the number of beds the home operated at the start of each year. The reliability of the measure is unknown because the consistency of counts was not reported. The assessment of validity is problematic because a clear definition of organizational size was lacking and no explanation was given regarding why the log number of beds was preferable to number of patients or employees, why the number at the start of the year was better than the end of the year or throughout the year, or why a log function was required.

When possible, the reliability and factor structure of CV measures should be examined to establish that the items adequately reflect the intended concept and that measurement error is not excessive. Researchers should also consider evidence for convergent and discriminant validity. Preferably, researchers will assess and correct measurement error using SEM or similar techniques. These steps are crucial because using poor measures of CVs can do more harm than good. For instance, if the measure used to represent a CV does not adequately reflect the intended construct, then the partialled relationships yielded by the measure cannot be interpreted as intended. Moreover, when CVs are measured with error, the parameter estimates for other variables in the analysis can be biased either upward or downward, leading to unwarranted conclusions (see Edwards, 2008, for the mathematical issues involved).

6. If the hypotheses do not include CVs, do not include CVs in the analysis.

This recommendation is a corollary of Recommendation 3 and is important in its own right because it encourages appropriate hypothesis testing and model specification. Just as good research practice requires strong logical connections between theory and hypotheses, it also requires correspondence between hypotheses and analyses. In other words, researchers should match analyses to their hypotheses. Assuring alignment between hypotheses and analyses means using only those CVs that are included in the statement of the hypotheses (Becker, 2005; Edwards, 2008). Too often, authors state hypotheses in terms of bivariate relationships between substantive variables but then include CVs in related analyses. In such cases, it is unclear whether the hypotheses as stated were supported. Further, as we emphasized earlier, the process of partialling created by statistical control results in coefficients that may not reflect the meaning of the original IVs. Thus, including CVs in analyses when hypotheses do not include the CVs results in invalid tests of the hypotheses. More broadly, including CVs in models clarifies how the CVs should be treated in the analyses. Is the CV proposed as a correlate of all IVs and cause of all DVs, or a subset of these? Or are some other kinds of relationships expected? We believe answering these questions will result in more thoughtful and meaningful analyses of CVs.

A related issue is that although CVs may not be the focus of a given study, they deserve the same analytical attention as do the variables of interest. For instance, if a model includes CVs, then the distribution of scores for the CVs needs to be as carefully examined as the distributions for the IVs and DV. Researchers should check for outliers in the CV because an extreme score can artificially inflate or deflate the correlation between an IV and CV, which, in turn, would distort the relation between the IV and DV. In addition, in many cases, researchers should examine possible nonlinear associations between the CV and IVs and potential CV × IV interactions. For example, a study in a top management journal examined the effect of leaders' holistic thinking (paying attention to the whole rather than the parts) on their behavior, controlling for length of experience as a leader. It seems possible that, like many learning curves, the impact of leader experience levels off as leader behaviors are more fully acquired. Also, a moderator effect might be plausible in that the effect of holistic thinking on leader behavior could be stronger when a leader has little experience than when he or she has a great deal of experience. Perhaps with substantial experience, leaders come to behave more automatically and rely less on active modes of thinking. Incorporating potential nonlinear and interactive effects of CVs such as these could substantively change the findings and conclusions of a study.

When available theory addresses connections between the CVs and other variables, the theory can serve as a guide to these kinds of analyses. Without a defensible guide, the analyses including CVs become problematic because there are numerous possible associations between CVs and other variables. For instance, there could be two- or three-way interactions involving the CVs and one or more IVs, or there could be quadratic or cubic relationship between a CV and other variables. These complexities are important to consider because if a researcher mistakenly assumes that relations involving the CV are linear or that the CV has only main effects, the results could be misleading. Researchers acting on Recommendations 1 and 2 are more likely to avoid misspecifications than are those who do not because these recommendations encourage the careful selection and consideration of the role of CVs.

7. Conduct comparative tests of relationships between the IVs and CVs.

We concur with Spector and Brannick (2011) that conducting appropriate comparative tests can help researchers understand the role of CVs and the relationships among other variables. For example, researchers interested in the relationship between employee commitment and organizational citizenship behavior (OCB) could be concerned that perceived supervisor support, a potential CV, has in some way influenced the observed relationship. As mentioned earlier, the CV might affect the measured but not latent variables (contamination), or it can affect the latent variables (spuriousness). Contamination occurs when extraneous (unintended) variables influence people's responses to scale items, hence biasing assessment. Such "nuisance" variables can be sources of common method variance that distort observed relations among variables (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). For instance, perhaps employees strongly supported by their supervisors want to please their bosses and, as a result, provide more socially desirable responses to organizational surveys. If so, supervisor support could artificially inflate the correlation between commitment and OCB scores. Supervisor support could affect the measure of commitment, measure of OCB, or both, and including supervisor support in analyses would have different effects depending on whether it contaminates one, the other, or both of the other measured variables.

Spuriousness exists when two latent variables are correlated because they share a common cause, not because there is a causal relationship between them. This would be the case if commitment and OCB were only related because perceived supervisor support was the cause of both. Researchers wanting to rule out spurious relationships might find including one or more CVs helpful. That is, if the commitment–OCB relationship remains after perceived support is statistically controlled, the link between commitment and OCB is unlikely to be entirely spurious. However, researchers cannot rely on statistical control alone to demonstrate a causal relationship. Ruling out spuriousness requires experimental designs that allow meaningful inferences of causality where it can be shown that a CV has no effect, or only a negligible effect, on the variables of interest. For instance, a researcher might create experimental conditions of low and high supervisory support to determine whether the link between commitment and

DOI: 10.1002/job

OCB varies across groups. A significant difference would indicate spuriousness, whereas, assuming sufficient sample size, a non-significant difference would argue against spuriousness.

If a researcher misspecifies the role of a CV in a hypothesis or model, the CV can produce misinterpretations even if it does not create contamination or spuriousness (Meehl, 1971). Researchers can avoid misconstruing results by generating alternative hypotheses regarding the CVs when planning a study. These hypotheses might specify whether the CV affects or is affected by other measured or latent variables (as in mediation or moderation) and whether contamination or spuriousness exist. Researchers have multiple methods at their disposal when conducting these analyses, including hierarchical regression and SEM. Spector and Brannick (2011) describe an approach whereby a researcher creates a set of baseline hypotheses regarding variables of interest and constructs alternative explanations involving the role of CVs. The researcher then conducts a series of analyses to evaluate the alternatives. We believe this kind of approach should become routine whenever researchers do not have strong arguments for the role of CVs in hypotheses and models.

8. Run results with and without the CVs and contrast the findings.

We strongly encourage researchers to run analyses with and without CVs because comparing the results demonstrates the impact of the CVs on the relations between the IVs and the DV (Atinc et al., 2012; Becker, 2005). Becker (2005) suggested that if the results do not differ, then only the analyses without controls need be reported, along with a statement that the results were essentially identical when CVs were included. It is possible to test the significance of the differences between corresponding effect sizes in the two analyses, but the process is complicated (Clogg, Petkova, & Haritou, 1995). We suggest that a practical decision rule is that if the standardized coefficients of the IVs with and without CVs differ by less than 0.1, then, in most cases, the differences are negligible.

A final point about running analyses with CVs is that the order of entry should be based on the goal of the analyses. In hierarchical regression, the convention is to enter CVs at Step 1 and enter IVs at one or more subsequent steps. If the CVs are theoretically meaningful and the focus is on incremental variance, then the conventional approach is sensible. However, it does not make sense to enter more tentative CVs (those with some supporting evidence but without a strong theory-based reason for inclusion) prior to the IVs. For instance, a researcher studying the connection between perceived injustice and employee theft might wish to control for workgroup norms. If there is a sound reason to control group norms (e.g., norms for deviant behaviors are known to exist and can influence theft through socialization), entering scores on a measure of norms at the first step of hierarchical regression could be reasonable.

The researcher may also want to control for impulsivity and contrary to Recommendation 2 might use number of prior jobs as a proxy, the notion being that more impulsive people tend to change jobs more often. We suspect that the basis for using number of jobs as a proxy is weak because people change jobs for reasons other than impulsiveness. If so, entering number of past jobs at Step 1 of a hierarchical analysis would be questionable. Although this CV is not necessarily atheoretical, the rationale is debatable. We concur with Carlson and Wu (2012) that IVs and theoretically meaningful CVs should be given the first opportunity to explain variance in outcomes because these variables offer important explanations for variance in the DV.

Finally, a researcher may expect that one or more CVs interacts with one or more IVs or has nonlinear relations with one or more DVs. To the degree that there are strong theory-based reasons to expect such relationships, they should be included in the hypotheses. If the researcher's expectations are more tentative, then interactive or nonlinear effects could be included in the last step of a hierarchical analysis. One problem with including many CVs in a study is that the number of possible interactions and nonlinear effects can become very large. We consider this another reason to take Recommendations 1 and 2 to heart.

Reporting and interpreting results

9. Report standard descriptive statistics and correlations for CVs, and the correlations between the measured predictors and their partialled counterparts.

DOI: 10.1002/job

This recommendation, if followed, will facilitate understanding of the psychometric properties of CVs, enhance potential for replication, and enable comparisons between measured and partialled predictors. Continuous and dichotomous CVs should be treated the same as other predictors with respect to reporting descriptive statistics, indices of reliability, and intercorrelations (Becker, 2005; Carlson & Wu, 2012). In addition, the amount of shared variance between the measured and partialled predictors should be reported because it has implications for inferences about hypotheses and external validity (Breaugh, 2008).

When using CVs, it is important to discuss the meaning of the scores on the residual predictors used to test hypotheses. This recommendation pertains to the issue of construct validity mentioned earlier in that as the shared variance between the original and partialled measures of the IVs declines, the partialled scores become less valid indicators of the IVs as originally conceived. If the CVs were incorporated into hypotheses, the researcher can interpret results in the context of the residual predictor, taking into account the size and direction of the effects and its level of statistical and practical significance. Regrettably, it is common in organizational research for a hypothesis to be stated in terms of a simple bivariate relationship between a predictor and a criterion but tested using a technique that removes the effects of other variables from the original predictor (Atinc et al., 2012). For example, when testing hypotheses via multiple regression, researchers often focus on regression weights rather than simple correlations even when correlations are more consistent with the nature of the hypotheses (O'Neill, McLarnon, Schneider, & Gardner, 2014). As previously discussed, this approach is inappropriate because including CVs can markedly change the substantive meaning of the constructs of interest. The discussion of results should align with the hypotheses or model and analyses.

An example is Breaugh's (2008) reanalysis of a study purporting to examine the effects of employee height on future earnings. Before examining this relation, the authors used multiple regression to remove the effects of gender, age, and weight. The relation between their residual height predictor and earnings was small ($\Delta R^2 = .02$). In comparison, the squared simple correlation between actual height and earnings was .10. These dissimilar effect sizes could result in different conclusions. Further, only 40% of measured height was reflected in the residual height predictor. In essence, height as reported in this study does not represent height as typically measured. Rather, the residual height predictor captures a different construct, perhaps physical proportionality. The distinction between height and physical proportionality could matter because, as noted by Breaugh, "…it is quite different to conclude that taller individuals earn more than it is to conclude physically well-proportioned individuals earn more" (p. 287). Employees have little control over their height, but they do have some control over their shape.

Consistent with Recommendation 8, researchers should also discuss differences between results including the CVs and those without. If there are no significant differences, researchers can note that the inclusion of CVs had little effect and perhaps suggest that the CVs may not merit future attention. If the results differ, then the results of both analyses deserve attention. Again, the interpretation will be much easier if the hypotheses and analyses are aligned.

10. Be cautious when generalizing results involving residual variables.

Restraint in extending findings based on partialling would improve assessments of external validity and practical applications of results. There are many examples of researchers inaccurately generalizing results based on residual predictors to a real world context. Meehl (1970) summarized the problem of statistical control as it relates to generalizability as follows: "In multivariate analysis, [the researcher] concocts statistically, by making certain algebraic 'corrections,' a virtual or idealized sample, the members of which are fictional people assigned fictional scores" (p. 401). That is, by using CVs, researchers create a statistical sample of individuals for which predictors that are correlated are forced to be orthogonal. We advise against inferring how phenomena operate in the population if variables that are correlated in the population are rendered uncorrelated through partialling.

Conclusion

Research in organizational settings is challenging owing to the difficulties of using experimental methods to study relevant phenomena. Therefore, researchers using correlational methods often rely on statistical control to rule out

DOI: 10.1002/job

the effects of extraneous variables, provide more accurate estimates of relationships among variables, or produce conservative tests of hypotheses. Unfortunately, the potential complexity of relations among variables, the uncertainty about why variables are related, and the possibility of overlooking important variables make it impossible to obtain these goals by merely including CVs in a study. The proper use of CVs requires careful selection based on theoretical considerations, a thorough and thoughtful analysis, and cautious interpretation of results.

Author biographies

Thomas E. Becker is a professor of management in the College of Business at the University of South Florida Sarasota–Manatee. He conducts research on employee commitment, motivation, performance, and organizational research methods and is a member of the Academy of Management and the Society for Industrial and Organizational Psychology.

Guclu Atinc is an Assistant Professor of Management at Texas A&M University–Commerce. His current research interests are in the area of corporate governance, such as upper echelons and board structure, along with research methods in strategic management. He is a member of Academy of Management and Southwest Academy of Management.

James A. Breaugh is a professor of management at the University of Missouri–St. Louis. He conducts research on the topics of recruitment, selection, work–family balance, work scheduling, and turnover. He is a fellow of the Society for Industrial and Organizational Psychology and the American Psychological Association.

Kevin D. Carlson is Professor and Head of the Department of Management in the Pamplin College of Business at Virginia Tech. Dr. Carlson's research focuses on the refinement of research methods to promote more systematic research progress. His other research interests include workforce analysis, staffing effectiveness, and improving individual productivity.

Jeffrey R. Edwards is the Belk Distinguished Professor of Organizational Behavior at the Kenan-Flagler Business School, University of North Carolina. His research addresses person–environment fit, polynomial regression, construct validation, structural equation modeling, and the development and evaluation of theory. He is a Fellow of SIOP and the Academy of Management.

Paul E. Spector is distinguished professor of I/O psychology, and director of the Occupational Health Psychology doctoral program at the University of South Florida. He is Point/Counterpoint editor for *Journal of Organizational Behavior* and Associate Editor for *Work & Stress* and is on the editorial board of *Journal of Applied Psychology*.

References

Atinc, G., Simmering, M. J., & Kroll, M. J. (2012). Control variable use and reporting in macro and micro management research. *Organizational Research Methods, 15,* 57–74.

Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods, 8,* 274–289.

Bernerth, J. B., & Aguinis, H. (2015). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*. DOI: 10.1111/peps.12103.

Bono, J. E., & McNamara, G. (2011). From the editors, publishing in AMJ—Part 2: Research design. *Academy of Management Journal, 54,* 657–660.

Breaugh, J. A. (2008). Important considerations in using statistical procedures to control for nuisance variables in non-experimental studies. *Human Resource Management Review, 18,* 282–293.

Burks, B. (1926). On the inadequacy of the partial and multiple correlation technique. *Journal of Educational Psychology, 17,* 625–630.

Carlson, K. D., & Herdman, A. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods, 15,* 17–32.

DOI: 10.1002/job

Carlson, K. D., & Wu, J. (2012). The illusion of statistical control: Control variable practice in management research. *Organizational Research Methods, 15,* 413–435.
Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology, 100,* 1261–1293.
Edwards, J. R. (2008). To prosper, organizational psychology should … overcome methodological barriers to progress. *Journal of Organizational Behavior, 29,* 469–491.
Loi, R., Yang, J., & Dienfondorff, J. M. (2009). Four-factor justice and daily job satisfaction: A multilevel investigation. *Journal of Applied Psychology, 94,* 770–781.
Meehl, P. E. (1970). Nuisance variables and the ex post facto design. In M. Radner, & S. Winokur (Eds.), *Minnesota studies in the philosophy of science, 4* (pp. 372–402). Minneapolis: University of Minnesota Press.
Meehl, P. E. (1971). High school yearbooks: A reply to Schwarz. *Journal of Abnormal Psychology, 77,* 143–148.
Nimon, K. F., & Oswald, F. L. (2013). Understanding the results of multiple linear regression: Beyond standardized regression coefficients. *Organizational Research Methods, 16,* 650–674.
O'Neill, T. A., McLarnon, M. J. W., Schneider, T. J., & Gardner, R. C. (2014). Current misuses of multiple regression for investigating bivariate hypotheses: An example from the organizational domain. *Behavior Research Methods, 46,* 798–807.
Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd edn). New York: Holt.
Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88,* 879–903.
Schjoedt, L., & Bird, B. (2014). Control variables: Use, misuse and recommended use. In Carsrud & M. Brännback (Eds.), *Handbook of research methods and applications in entrepreneurship and small business* (pp. 136–155). Northhampton, MA: Elgar Publishing.
Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods, 14,* 287–305.
Spector, P. E., Zapf, D., Chen, P. Y., & Frese, M. (2000). Why negative affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior, 21,* 79–95.
Williams, L. J., Vandenberg, R. J., & Edwards, J. R. (2009). Structural equation modeling in management research: A guide for improved analysis. *The Academy of Management Annals, 3,* 543–604.
Wright, T. A. and Bonett, D. G. (2007). Job satisfaction and psychological well-being as nonadditive predictors of workplace turnover. *Journal of Management, 33,* 141–160.

DOI: 10.1002/job